Ingénierie des Systèmes d'Information
Vol. 30, No. 6, June, 2025, pp. 1629-1638
Journal homepage: http://iieta.org/journals/isi

Enhancing Fake News Detection via PSO-Optimized Ensemble Learning: A Comparative Study of SVM, NB, and RF

Zainab Khyioon Abdalrdha[1*], Amal Abbas Kadhim[2], Amaal Kadum[2], Wedad Abdul Khuder Naser[2]

[1] Department of Computer Science, College of Basic Education, Mustansiriyah University, Baghdad 10001, Iraq
[2] Department of Computer Science, College of Education, Mustansiriyah University, Baghdad 10001, Iraq

Corresponding Author Email: zainabkhyioon83@uomustansiriyah.edu.iq

https://doi.org/10.18280/isi.300621

Received: 17 May 2025
Revised: 18 June 2025
Accepted: 24 June 2025
Available online: 30 June 2025

Keywords:
fake news detection, ensemble learning, Particle Swarm Optimization (PSO), TF-IDF, machine learning, Support Vector Machine (SVM), Naive Bayes, Random Forest (RF)

ABSTRACT

Given the rapid spread of fake news across digital platforms, there is a pressing need for a reliable and efficient detection method. Current ensemble learning models often lack optimal weight tuning, limiting their performance in fake news classification tasks. To address this gap, we propose a Particle Swarm Optimization (PSO)-optimized ensemble model that integrates Support Vector Machines (SVM), Naive Bayes, and Random Forest (RF) classifiers using a soft voting strategy. Text data is preprocessed and transformed into numerical features using TF-IDF vectorization. The dataset, derived from the ISOT Fake News corpus, is split into training (80%) and testing (20%) subsets. Each base classifier is individually trained and evaluated, followed by the construction and assessment of an unoptimized voting ensemble. Subsequently, PSO is employed to fine-tune the weights of the base classifiers within the voting ensemble, enhancing overall prediction performance. The optimized model achieves 98.32% accuracy and an F1-score of 98.33%, outperforming both the unoptimized ensemble and standalone classifiers, as well as surpassing several state-of-the-art methods. This approach not only improves detection accuracy but also offers a scalable, interpretable, and effective solution to the fake news problem. Performance is evaluated using standard metrics such as ROC curves and confusion matrices, providing a comprehensive assessment of the model's reliability.

1. INTRODUCTION

Digital platforms have greatly improved communication, but they also make fake news more likely to proliferate, a major problem for public health, political stability, and societal confidence [1]. The expansion of misleading information-spreading social media channels has contributed to the worldwide problem of fake news—defined as intentionally false content passing for real news [2]. Among other things, fake news influences elections, fuels violence, and undermines public confidence in trustworthy institutions [3]. Therefore, efficient automated solutions to stop the dissemination of fake information are much sought for. One-model ML methods such as LR or SVM typically form the basis of conventional approaches to spotting false news [4]. The complex and multi-dimensional structure of disinformation may be challenging for them to understand. New ensemble learning methods that incorporate the best features of several models have shown promise; nevertheless, to improve detection accuracy [5]. Since ensemble learning exploits the diversity of base classifiers to improve generalization and robustness [6], it is useful in applications including fake news detection, where data is often imbalanced and noisy. Still, there are challenges even with these advances. Many present methods fall short of the potential of ensemble learning due to inadequate feature engineering and suboptimal model integration [7]. Models also have to be constantly refined and improved to fit the often-shifting patterns of false information disseminated by fake news [8]. To address these issues, our work offers an enhanced ensemble learning approach combining modern ML models with advanced feature extraction techniques and hyperparameter modification. This paper makes two significant contributions: The first is an ensemble learning framework combining SVM, NB, and RF using a voting classifier; the second is a PSO optimization leveraging base classifier weight optimization to improve ensemble performance. This paper is organized generally as follows: Section 2: An introduction to the literature review on fake news identification and ensemble learning. Section 3 covers the advised methodology. Section 4 deals with the experimental design and data and investigates the outcomes; Section 5. Section 6 provides conclusions and directions for further research on the work.

2. LITERATURE REVIEW

Fake news detection has garnered significant attention due to its profound implications on public health, political discourse, and societal trust. In response, numerous studies have leveraged machine learning (ML) and deep learning (DL) algorithms to address the challenge of identifying deceptive

content across various platforms and languages. Early approaches explored traditional ML classifiers such as Support Vector Machines (SVM), Naive Bayes (NB), Random Forests (RF), and Logistic Regression (LR), often paired with feature extraction methods like TF-IDF, n-grams, and sentiment-based analysis. For instance, one study [9] utilized a combination of lexical features, sentiment analysis, and embeddings (GloVe and character-based) alongside TF-IDF, comparing models including SVM, LR, Decision Trees (DT), LSTM, convolutional HAN, and character-level CLSTM. The results showed that an LSTM model trained with NB and bigram TF-IDF features reached a peak accuracy of 94%.

More advanced architectures have shifted towards analyzing the emotional flow of articles. The Fake News Flow model [10], designed to detect emotional manipulation in longer texts, implemented CNN/Bi-GRU neural architectures, outperforming other DL models such as LSTM, HAN, BERT, and Longformer with an accuracy of 96%, a recall of 97%, and a macro F1-score of 96%. This highlights the effectiveness of modeling emotional content flow in detecting fake narratives. In another line of research, social context has been integrated with textual features to enhance performance. One study [11] combined entropy-based feature selection and min-max normalization with stacked classifiers (SVM, RNN, RF), achieving 81.9% accuracy, significantly reducing false positives. Similarly, FNC-1—a widely used English dataset—was used to evaluate bidirectional LSTM models and multi-head LSTM models, yielding competitive accuracy scores of 85.3% and 82.9%, respectively [12]. N-gram analysis continues to play a central role in textual analysis, particularly when combined with ML classifiers. Research has demonstrated that combining TF-IDF features with classifiers such as SVM and LR can achieve precision rates up to 92% [13]. Another hybrid approach [14] included both news article content and user comments, processed through RNNs, GRUs, and SVMs, showing improved robustness in fake news classification through multi-source data integration.

Ensemble learning has also emerged as a powerful strategy to boost classification performance. A study targeting Malay-language news [15] employed a hybrid ensemble of various ML techniques, achieving significant improvements across accuracy, precision, recall, and F1-score. Meanwhile, sentiment analysis and metadata features were applied to Arabic news, where the best-performing models (e.g., DT, AdaBoost, LR, RF) achieved up to 76% accuracy [16]. Several efforts have focused specifically on Arabic fake news detection. One study [17] proposed an ensemble approach combining TF and TF-IDF feature extraction with XGBoost, CatBoost, and NGBoost classifiers. This methodology enhanced classification accuracy and reduced false positives. In a similar vein, Alkhair et al. [18] constructed a novel Arabic corpus from YouTube comments and applied DT, SVM, and Multinomial NB to analyze rumor propagation, achieving an SVM accuracy of 95.35%. Suhasini et al. [19] introduced a hybrid DL framework that integrated CNN and LSTM with traditional classifiers such as SVM, NB, KNN, and LR. The CNN-LSTM-SVM combination yielded the highest accuracy at 96%, demonstrating the strength of merging feature extraction from DL with the decision boundaries of ML classifiers. The use of vectorization methods remains central in recent works. Thaher et al. [20] tested various ML techniques to identify optimal text representation strategies, finding that TF-based models outperformed LR with an accuracy of 82% and an F1-score of 80.42%. Likewise, ensemble learning combined with parameter tuning, as explored in [21], demonstrated that stacking and delegation methods using TF-IDF and count vectorizer preprocessing could surpass individual models in both AUC and F1 metrics. A related study [22] employed RapidMiner and Python to preprocess Arabic comments and tested multiple ML classifiers, with NB achieving 87.18% accuracy. Further enhancement was reported in a soft-voting ensemble model combining NB, SVM, LR, and RF, optimized via Grid Search CV, achieving 93% accuracy on the Kaggle dataset [23].

Beyond traditional models, transformer-based architectures have also been explored. The study [24] extended the scope to include BERT Multilingual, RoBERTa, and ALBERT for Indonesian fake news detection. Using contextual embeddings from deep Transformer layers, ALBERT emerged as the best performer with an accuracy of 87.6%, surpassing the others in both precision and F1-score. Finally, an advanced approach [25] combined CNN and BiLSTM in a PSO-optimized hybrid model, applied to the LIAR dataset with GloVe and FastText embeddings. This method achieved an impressive 96.8% accuracy, outperforming both traditional and Transformer-based models. The model demonstrated high robustness, strong generalization, and effective hyperparameter tuning. Collectively, these studies reveal a progressive shift from traditional feature-based ML methods to hybrid and ensemble DL architectures, particularly those optimized through metaheuristic algorithms. They underscore the importance of combining diverse data sources—text, sentiment, context, and user interactions—for achieving robust fake news detection systems.

3. SUGGESTED APPROACH

The suggested approach described in this section employs ensemble learning with TF-IDF for feature extraction and (PSO) to optimize model performance for fake news detection, as shown in Figure 1, the PSO-Ensemble Model Fake News Architecture:

This suggested approach searches social media for fake news using a six-step procedure. The relative position of a news article's headline-oriented assessment motivated this research.

1- Preprocessing the data set to convert unstructured data sets into structured ones comes first in the approach.
2- In the second stage, TF-IDF feature extraction is employed to identify undiscovered fake news traits and different correlations between news articles.
3- The study employs an ensemble learning approach based on base ML models SVM, NB, and RF with TF-IDF for feature extraction to enhance fake news classification.
4- Voting Classifier (Unoptimized) – Aggregates predictions from base models using soft voting.
5- The PSO-Optimized Ensemble Model uses PSO to find the best weight combination for SVM, NB, and RF, which improves classification accuracy. For a more realistic setup, we constrained the Random Forest to 30 trees with a maximum depth of 5, utilized a Radial Basis Function (RBF) kernel with a value of C=0.002 in the Support Vector Machine, and set the high variance smoothing parameter in the Naive Bayes to 30 and using PSO with a swarm size of 10 and 20 iterations.

6- Model Evaluation: measure how well the fake news detection model performs; several key evaluation metrics are used.

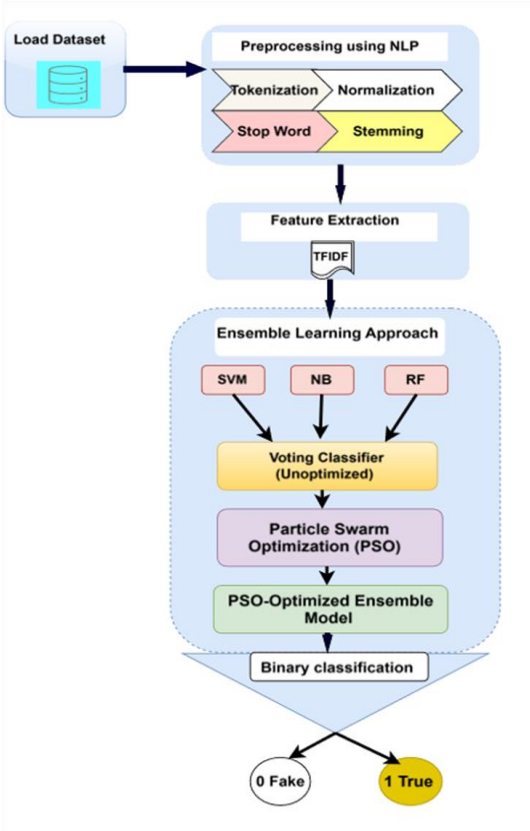


Figure 1. Fake news architectural PSO-ensemble model

Ultimately, our study effectively identified fake news. Every one of these phases will be covered in more detail in the next subsections.

A. Data collection

Kaggle provided ISOT Fake News [26, 27]. The dataset includes fake and real news. These are Reuters.com stories. Much information was inaccurate. Politifact and Wikipedia listed bogus news websites. Foreign and political news predominate. Two CSVs contain data. "True.csv," the first file, contains 21,417 Reuters.com items. The second file, "Fake.csv," has 23,481 fake news items. Titles, bodies, categories, and dates define articles. The fake news dataset from Kaggle.com was matched with 2016–2017 stories. Figures 2 and 3 show articles and categories, and fake news type distributions.

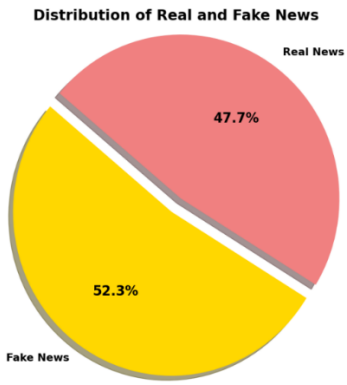


Figure 2. The summary of the ISOT fake news dataset

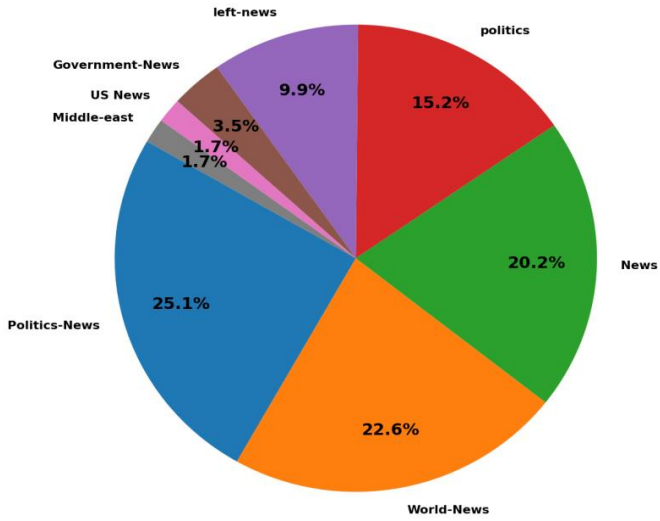


Figure 3. Distributions of fake news types

B. Data preprocessing

NLP processes and analyzes vast quantities of speech and text within the realm of AI. The primary objective is to structure unorganized text for analysis. Natural Language Processing employs techniques for data processing and organization. This research employed tokenization, linguistic components (normalization, stop word elimination, and stemming), and text vectorization. Tokenization, the process of segmenting text into terms or words, is fundamental in NLP. Stop words, prevalent terms may be excluded without sacrificing meaning. Consistency necessitates normalization. Text normalization standardizes non-standard lexicon. Normalization standards are implemented in the English language. Normalization enhances consistency by standardizing text, such as converting all characters to lowercase, removing extra spaces, and correcting common misspellings. Stemming, a heuristic technique for normalization, reduces inflected or derived words to their root form; for example, "running," "runs," and "ran" are all reduced to "run." The text's language influences the choice of stemming algorithms, with the Porter Stemmer being one of the most widely used for English [28, 29].

C. Feature extraction

Feature extraction or text vectorization extracts numerical features from unstructured text. Knowledge extraction is machine learning's forte. As a multivariate sample or vector, word-based statistical measures can provide numerical text attributes [30]. Word occurrences are weighted in this model. Many term weighting methods use TF-IDF and TF [31, 32]. The optimal text representation model for false news identification is determined using both methods. TF weight depends on the frequency of each phrase 't' in a text. Word count vectors illustrate text. Average phrases are unimportant. This problem is solved by employing the logarithm function in Eq. (1) and other normalization-based TF methods:

$$W_t = \begin{cases} 1 + log_{10} TF_t & TF_t > 0 \\ 0 & Otherwise \end{cases} \quad (1)$$

where, 'TFt' represents the frequency of natural terms and 'wt' represents the frequency of weighted terms in the text. BTF is an additional text format that relies on TF. It takes text and turns it into a binary vector that shows whether words are there or not. No phrase is given more weight than any other using

traditional frequency-based weighting, even though fewer common phrases frequently contain more information. We need to adjust our metric such that common phrases are given more weight and less weight than unusual ones. We used DF, IDF, and TF-IDF as our word weighting strategies. Eq. (2) can be used to generate IDF to quantify word frequency, where DF is the number of corpus documents (texts) that include a phrase:

$$IDF_t = log \frac{N}{DF_t} \quad (2)$$

A high score is given to odd phrases that appear in multiple papers, where 'N' reflects the number of documents in the collection. To make the most of both the TF and IDF measurements, an effective statistical weighting method called TF-IDF is suggested for word weighting in text categorization and information retrieval (Eq. (3)). Therefore, a term is given a lot of weight in the text if it exists in a few articles, but a lesser score if it appears in the majority or just a few documents:

$$TF - IDF_t = W_t \times IDF_t \quad (3)$$

To determine 'Wt' and 'IDFt,' we utilize Eqns. (1) and (2), respectively.

D. Ensemble classifiers

The ensemble model integrates three base classifiers, each chosen for its unique strengths:

1. SVM:

This algorithm can solve regression and classification problems when supervised. Classification issues: Use it often. SVMs divide data into regions, making them powerful ML classifiers. The SVMs strive to identify the largest margin that splits the dataset in half, and then assign new data to one of the two groups. SVMs are popular for their accuracy and inexpensive processing power. It excels with smaller datasets. SVMs can handle multidimensional spaces and are memory-efficient [33].

2. Naive Bayes algorithm

Naive Bayes estimates conditional probability, estimating whether an event will occur if another has already occurred. This classification strategy utilizes Bayes' Theorem and predictor independence. One feature in a class is independent of others in NB. Naive Bayes is fast, easy to implement, and effective for huge datasets. Text classification with binary and multiclass classifications is reliable [34].

3. RF classifier

RF is an adaptable, straightforward, and diverse supervised ML method. The challenges of classification and regression can be resolved by it. The forest it constructs is a collection of DT models working together to improve forecast accuracy. When it comes to categorization, each DT works independently to forecast a class's result; the one with the most votes at the end gets the last say [35].

4. Voting ensemble classifier

Voting ensemble classifiers use many models to improve classification accuracy and resilience. Hard voting involves all models voting for a class and the majority decides; soft voting involves averaging the expected probability; both are necessary for its operation. Taking advantage of model strengths enhances stability, reduces overfitting, and increases generalization. This tool is useful for classification problems in finance, healthcare, and NLP for risk assessment, sickness prediction, and sentiment analysis [36]. Based on a soft-voting ensemble classifier for SVM, NB, and RF, this paper.

5. PSO ensemble model

The PSO algorithm was created by Eberhart and Kennedy. It was inspired by birds' smart food-finding. As a swarm intelligence optimization algorithm, PSO is stable, converges quickly, has few parameters, and is easy to apply. PSO has been used to optimize data mining, artificial neural network training, vehicle path planning, medical diagnostics, and system and engineering design [37].

5.1.1 Basic PSO solution

Based on swarm social behaviors like fish in a school and birds in a flock, Kennedy and Eberhart created PSO, a population-based self-adaptive optimization method. The PSO method searches the objective function landscape by quasi-stochastically modifying particle paths.

Every particle follows its best experience and the swarm's global best solution to alter its velocity and position. Eqs. (1) and (2) [33] prescribe the update equations for the velocity V_i^{t+1} and location X_i^{t+1} of the ith particle at the dth dimension in the PSO model:

$$V_i^{t+1} = \omega * V_i^t + c_1 * r_1 * (P_{besti}^t - X_i^t) + c_2 * r_2 * (g_{besti}^t - X_i^t) \quad (4)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \quad (5)$$

where, v_i and x_i are the particle's velocity and position, and p_{besti} and g_{best} are its historical and global best solutions. Additionally, c_1 and c_2 are location constants, whereas r_1 and r_2 are random values from [0, 1]. Additionally, t and w reflect the current iteration number and inertia weight.

E. Model evaluation

Comparing text-based fake news detection systems requires performance evaluation criteria to determine classifier accuracy and efficacy. The applied experimental methodology uses multiple methodologies. An evaluation confusion matrix or contingency table was utilized. The contingency table covers TP, TN, FP, and FN donations. TP and TN contributions are excellent for classifying positive and negative situations. FP represents negative cases misclassified as positive, while FN represents positive cases misclassified as negative. Classifier performance formulae are in Table 1 [38-40].

Recall: The percentage of positive cases the model recovered.

Precision, also known as **Positive Predictive Value**, is a statistic used to evaluate classification algorithms, especially where positive case accuracy is important.

The **F1 Score** is used to evaluate classification models, especially when balancing Precision and Recall. When data is imbalanced or False Positives and Negatives are equal, it works best.

Accuracy is a typical classification model performance metric. It shows the percentage of model predictions that were right (positive and negative). The percentage of cases classified as True Positives and True Negatives out of the total number is shown.

Area under the ROC Curve (AUC) is a typical statistic for evaluating classification models, especially binary ones. It tests the model's ability to discriminate positive and negative classes at all threshold levels.

ROC: A graph for assessing binary classification models. The True Positive Rate (TPR) and False Positive Rate (FPR) change with the prediction threshold (Threshold) used by the

model to classify samples.

Table 1. Performance evaluation metrics

Metric Name	Formula
Sensitivity or Recall (R.)	$\frac{TP}{TP+FN}$
Precision (P.)	$\frac{TP}{TP+FP}$
F1 Score (F1.)	$2 * \frac{Precision*Recall}{Precision+Recall}$
Accuracy (Acc.)	$\frac{TP+TN}{TP+FN+TN+FP}$
Area Under ROC Curve (AUC)	$0 \leq Area\ under\ the\ ROC\ Curve \leq 1$
ROC	1 – specificity

4. EXPERIMENTAL RESULTS

This section talks about the results from testing SVM, Naive Bayes, RF, the unoptimized voting classifier, and the PSO-Optimized Ensemble Model for detecting fake news, following the earlier explained method. The analysis focuses on evaluating the performance of each model using Acc., P., R., and F1. Metrics based on the labeled dataset, which was applied by directly classifying each news article that was strictly labeled as "fake" or "real" with no specific topics in mind. The classification problem was reduced to a binary classification problem on article credibility, based on content features available from the ISOT dataset. Table 2 presents a discussion of the results, based on performance metrics.

1. Performance of different models

•**SVM:** Performed poorly with an accuracy. of only 54.45%, indicating that it is not suitable for classifying this dataset. This could be due to improper hyperparameter tuning or the model's incompatibility with the data structure.

•**Naive Bayes:** Showed moderate performance with an Acc.of 87.75%. While it works well with textual data, it might struggle with complex feature interactions.

•**RF:** Delivered a strong performance with an Acc. of 97.99%, suggesting that it effectively handles multiple features and identifies patterns efficiently.

•**Voting (Unoptimized):** Achieved an Acc. of 93.54%, slightly lower than RF but still a robust model benefiting from the ensemble effect.

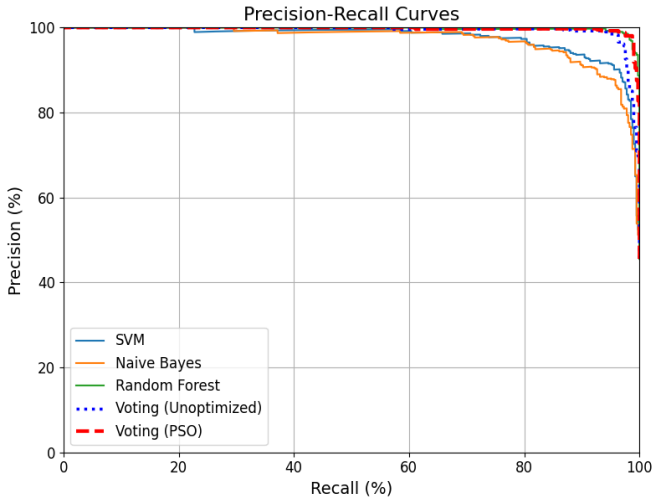
•**Voting (PSO-Optimized):** The best-performing model, with an Acc. of 98.32%, proving that PSO-based weight optimization significantly enhances performance.

Table 2. The results from the various models

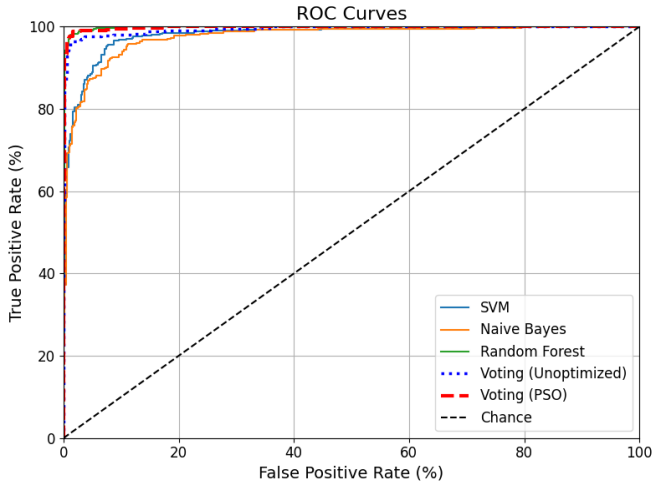
NO	Model	Acc.	p.	R.	F1.
0	SVM	54.454 %	75.19%	54.45 %	63.18 %
1	Naive Bayes	87.75%	89.24%	87.75 %	87.50 %
2	RF	97.99%	98.01%	97.99 %	97.99 %
3	Voting (Unoptimized)	93.54%	93.97%	93.54 %	93.55 %
4	PSO-Optimized Ensemble Model	98.32%	98.33%	98.32 %	98.33 %

2. Analysis of confusion matrices

This section will clarify the key Precision-Recall Curve and the ROC curve, related to the findings of the proposed technique, as shown in Figure 4. The model performance dashboard includes parts A and B.



(a)



(b)

Figure 4. (a) Model performance precision-recall curve part A; (b) Model performance dashboard part B

Summarized in the following points of Figure 4 in parts A Precision-Recall Curve and ROC Curve in parts B:

- Precision-Recall Curve Summary Shows the trade-off between precision and recall for different classification models. Models Compared: SVM (blue solid line), Naive Bayes (orange solid line), Random Forest (green solid line), Voting (Unoptimized) (blue dotted line), and Voting (PSO) (red dashed line).
- Best Performance: Voting (PSO) has the highest and most stable precision across the entire recall range. Closely followed by Random Forest and Voting (Unoptimized).
- Moderate Performance: Naive Bayes performs reasonably well but starts dropping in precision at higher recall levels.
- Lowest Performance: SVM shows the weakest performance, especially in high recall zones.
- Interpretation: Models closer to the top-right corner (precision ~100%, recall ~100%) are better.

Voting (PSO) outperforms all other models in both ROC and precision-recall evaluations, making it the most reliable classifier for your dataset.

Table 3. The experimental settings and model hyper parameters

Model	Parameter	Value	Optimized Weights (via PSO)
SVM	Max Iterations	1000	SVM weight (0.0316)
	Regularization Parameter (C)	0.002	
	Kernel Type	RBF (Radial Basis Function)	
Naive Bayes (Gaussians)	Variance Smoothing	30	Naive Bayes weight (0.4228)
Random Forest	Number of Trees	30	Random Forest weight (0.5456)
	Max Tree Depth	5	
PSO Optimization	Swarm Size	10	-
	Max Iterations	20	
	Weight Bounds	[0.001, 100]	

The ROC (Receiver Operating Characteristic) Summary shows the true positive rate and false positive rate of 5 models in Figure 4, part B.

- Model Comparisons: SVM – Blue solid line, Naive Bayes – Orange solid line, Random Forest – Green solid line, Voting (Unoptimized) – Blue dotted line, Voting (PSO) – Red dashed line, Chance line – Black dashed diagonal (baseline).
- Voting (PSO): Achieves the best performance and hugs closely the top-left corner.
- Indicates an extremely high true positive rate with very few false positives.
- It also confirms its strength at classification, again as evidenced by tabular metrics (Acc = 98.32%).
- Random Forest and voting (non-optimized): Both look pretty good following the PSO curve.
- Propose robust classifiers that have nearly similar performance to PSO.
- Naive Bayes: Not doing bad, but scoring a bit less than RF and voting models.
- It grows the fastest; however, it doesn't hug the top-left like the others.
- SVM: The worst-performing model of all the models.
- It is located further from the top left ideal corner and has high FPR and low TPR.
- Also in line with low reported accuracy (Acc = 54.45%).
- Random line (diagonal): It is also a line indicating random guessing.

Specific hyperparameters were chosen for each model after preliminary testing to ensure optimal performance. A summary of each model's hyperparameters is provided in Table 3. The experimental settings and model hyperparameters.

3. Discussion of the drawbacks and advantages of the proposed method

Several major limitations of existing fake news detection approaches were identified in earlier studies. Dataset bias, insufficient generalization, and poor short text performance plague traditional and sophisticated false news detection methods [9]. High-level models are impossible to interpret, limiting their use in healthcare and journalism. Fake flow is popular [10], but its lexicon-based emotive features may not work across languages or fields. The emotional component needs feature fusion since it works poorly without topic information. They don't have reinforcement learning and rely on large, high-quality datasets to perform well [11]. Although accurate, the proposed models vary substantially amongst datasets, raising generalizability problems [12]. CNN and LSTM AutoEncoder struggle with complex data, requiring more flexible architectures. Models using basic n-gram features along with their linear classifiers respond to the problems of n-gram size sensitivity, shallow text representation and insufficient generalization [13]. Hybrid models, such as those that fuse content and user comments [14], were also susceptible to adversarial attacks, heavily dependent on the size of feature vectors, and restricted by binary classification techniques. Similar methods failed with small and unbalanced datasets [15, 17], focusing on lexical terms in naive feature extraction [15, 19], reviewing issues in language-specific challenges, such as the Arabic case [16, 18], and underutilizing deep learning with or without up-to-date semantic features [15, 19, 21]. Additionally, methods like blending or voting can improve accuracy but come with problems like being complicated, relying on simple features, the chance of overfitting, and lacking clarity or strength when faced with noisy or tricky data. In contrast, the PSO-Optimized Ensemble Model can deal with these drawbacks well. By simultaneously performing PSO optimization with a swarm size of 100 and coefficients calculated in 20000 iterations, the accuracy of the optimized ensemble system is 98.32%, much better than all the single models and the voting ensemble with no optimization. This fact shows the accuracy of PSO for each voting weight adjustment, even when several particles are small in the suggested method. Furthermore, it achieves a higher accuracy compared to its predecessors. Unlike approaches that relied on handcrafted features, an optimized ensemble can leverage on the flexibility of dynamically and strategically tuning the model at the input, making it robust against noisy inputs and less prone to adversarial attacks. Moreover, with the help of integrating multiple classifiers (e.g., SVM, NB, RF) by using the PSO-optimized soft voting mechanism, the system improves the diversity for the models, thereby reducing the over-reliance on a single model architecture and promoting the generalization ability across various forms of fake news. Accordingly, the PSO-Optimized Ensemble Model increases the predictive performance, resists the noise and outliers, adapts the changes and works in an efficient way, and eliminates the principal constraints found in previous literature.

There are still several restrictions in spite of these benefits. Because PSO is an offline process that is only done once, its computing cost can be somewhat high, particularly during optimization. Nevertheless, this makes it suitable for a wide range of real-world applications. Furthermore, TF-IDF characteristics offer a portable and comprehensible substitute that permits quick training on common hardware, even though their use may not be as semantically rich as transformer-based embeddings. Last but not least, our balanced sampling and preparation procedures assist in lessening the influence of potential dataset biases, even though they exist like any real-world data. All things considered, the PSO-Optimized Ensemble Model effectively overcomes important limitations identified in earlier research and provides a convincing

balance of accuracy, efficiency, and adaptability.

Summarizing the ten studies discussed in the Literature Review, focusing on the datasets used, algorithms, proposed methods, and best performance achieved in Table 4 comparison of fake news detection studies:

Table 4. Comparison of research on fake news detection

Reference	Dataset Used	Algorithms	Proposed Method	Best Performance
[9]	1. Politics LIAR (12.7k samples) 2. Fake or Real News (6.3k samples): 2016 US election 3. Politics, economy, health, etc. (79.5k)	SVM, LR, DT, LSTM, convolutional HAN, RoBERTa, and character-level CLSTM.	RoBERTa	Accuracy: 96%
[10]	700 and 500 datasets, in addition to one dataset that was developed by the author	A Convolutional Neural Network (CNN) and Bidirectional Gated Recurrent Units (Bi-GRUs)	CNN	Accuracy: 96%
[11]	PHEME dataset (103,212, including user comments and original content)	(SVM), (RNN), (RF); RF used as meta-classifier	Stack ensemble	Achieved 81.9% accuracy
[12]	FNC-1 comprises train bodies (1683 articles) and stances (49972 headlines).	Bidirectional LSTM concatenated and multihead LSTM models.	Concatenated FND_Bidirectional LSTM Model	Accuracy: 85.3%
[13]	- BuzzFeed's dataset comprises 12,600 false and 12,600 real news stories about the 2016 US elections.	ML Detection Methods: DT, L.R., SGD, and LSVM.	LSVM + LR	Accuracy: 94.5%
[14]	- Fake News net dataset, Politi Fact, Gossip Cop	Text & Comment Feature Extraction RNN with BGRUnits. - SVM with a Gaussian kernel for final classification.	RNN-GRU + SVM model	Accuracy: 91.2% Recall: 96.1% For PolitiFact
[15]	1,000 news articles Data collected between 2017 and 2020.	Using TF-IDF and Bag-of-Words (BoW) for FE techniques. Single Classifiers: SVM, LR, NB, DT, and KNN.	- Hybrid Ensemble Model - Voting-based ensemble approach. - Combines classifier strengths.	Accuracy: 75%
[16]	Text includes 1,822 Syrian tweets.	- The L.R., D.T., R.F., and Ada Boost algorithms are used in the study to analyze sentiment analysis. Objective: Create a binary classifier to identify tweets as 'untrustworthy' or 'trusted' using probability estimations.	Random Forest Ada Boost	Accuracy: 76% of R.F, 77% of Ada Boost
[17]	- Fake news samples: 1,158 articles - Real news samples: 1,380 articles	Feature Extraction using TF-IDF This study employs ensemble learning models to improve classification performance are XGBoost, Cat Boost, and NG Boost	Cat Boost (TF-IDF)	Accuracy: 91.1%
[18]	- YouTube Comment Data Analysis: Gathered from 4079 comments.	- Investigating Middle East Fake News -Utilized Arabic comments on YouTube. -Utilized MNB, DT, and SVM.	SVM	Accuracy: 95.35%
[19]	- Dataset from Kaggle. Total News Articles: 25,117	Deep Learning Models: CNN – Extracts features from text. LSTM – Captures long-term dependencies. DNN – Enhances feature learning. Machine Learning Classifiers (Ensemble Model): SVM, NB, KNN, LR, and SoftMax Classifier	CNN-LSTM-SVM	Accuracy: 96%, and Recall:97%
[20]	1862 Arabic Twitter	Research on Fake News in Arabic Tweets - Uses NLP, ML, Harris Hawks Optimizer for feature selection. -Model includes K-NN, RF, SVM, NB, LR, DT, XGBoost.	L.R. classifier performed the best	Accuracy: 82%
[21]	- Source: Kaggle Fake News Dataset - Size: 20,000+ articles	TF-IDF and Count Vectorization for feature extraction ML Models: using [LR, SVM,	Stacking: probability-based. Delegation (Iterated)	Accuracy: 96.94% for Stacking:

		XGBoost DL Models: using LSTM) and CNN Ensemble Learning Models: Stacking (Stacking Label and probability-based). Delegation (Fall and Iterated) Hyperparameter Tuning: using Grid Search & Random Search.		probability-based. 98.15% for Delegation (Iterated)
[22]	- Create their dataset.	-Identifying Arabic Fake News in Social Media Comments - Utilizing KNN, DT, SVM, NB.	SVM	Accuracy: 87.18%
[23]	Kaggle Fake News Dataset [Size: 6,335 news articles]	Use Count Vectorizer for feature extraction. Hyperparameter Tuning: Used Research to find the best parameters for each model. The study uses four ML models NB, SVM, LR, and RF for Soft Voting Ensemble.	Soft Voting Ensemble	Accuracy: 93%, Precision: 94%, F1-Score: 93, 93%
[24]	Three datasets from the following sources"Indonesian hoax news detection", turn back hoax-dataset (GitHub),and Hoax-NewsClassification (GitHub) were merged into a single dataset for the paper.	- Investigation into Transformer-Based Models The use of contextualized vector representations is employed. Using previously trained models, it creates an embedding. - A comparison was made between four Transformers models: ALBERT, RoBERTa (Indonesian version), IndoBERT, and BERT-Multilingual	ALBERT	Accuracy: 87.6%, Precision and F1-Score: 86.9%.
[25]	LIAR dataset	Combines BiLSTM and CNN BiLSTM: Captures sequential dependencies in text CNN: Extracts spatial features PSO: Optimizes key hyperparameters, including: Learning rate, Batch size, Number of CNN filters, Number of LSTM hidden units Text Representation: Uses GloVe and Fast Text word embeddings	BiLSTM + CNN hybrid model optimized using PSO.	Accuracy: 96.8% F1-Score and Precision: both exceeded 95%
The proposed method	Fake and Real News Articles	(RF), (NB), and (SVM), Voting (Unoptimized), and PSO-Optimized Ensemble Model	PSO-Optimized Ensemble Model	Accuracy: 98.32%

5. CONCLUSIONS

In this paper, we presented an ensemble learning model for fake news classification. Following an initial pre-processing, the data was processed to train the various ML such as SVM, NB, and RF models separately. Two techniques were used to assemble the winners' strengths: a soft voting ensemble to amalgamate their strengths and the PSO method for fine-tuning the ensemble parameters. Results indicate that, in comparison with single models, ensemble learning substantially outperforms. The performance of SVM and NB was fair, but RF performed better. The PSO-Optimized Ensemble Model, however, weighed P., R., and F1., achieved the highest acc. of 98.32%. From the practical viewpoint, the experiment results show that efficient classifier tuning can improve the framework, and the optimization algorithm, such as PSO, still keeps great power in feature selection and classifier tuning for fake news detection. These findings verify the excellent function of ensemble learning with intelligent optimization for reliable and accurate fake news classification. Future research can explore integrating deep learning models such as transformers and LSTMs with ensemble learning to further improve fake news classification accuracy. Additionally, incorporating social network-based features, such as user credibility scores and engagement metrics, could enhance the robustness of the classification model. Optimizing ensemble models with advanced metaheuristic algorithms, such as genetic algorithms and differential evolution, may also improve performance. Lastly, expanding the dataset to include multilingual fake news sources will help generalize the model for broader applications.

ACKNOWLEDGMENT

The authors would like to thank AL_Mustansiriyah University (www.uomusiriyah.edu.iq), Baghdad-Iraq for its support in the present work.

REFERENCES

[1] Shu, K., Cui, L., Wang, S., Lee, D., Liu, H. (2020).

dEFEND: Explainable fake news detection. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 395-405. https://doi.org/10.1145/3292500.3330935

[2] Tacchini, E., Ballarin, G., Vedova, M.L.D., Moret, S., de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. Journal of Computational Social Science, 3(1): 1-20.

[3] Tucker, J.A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3144139

[4] Bondielli, A., Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. Information Sciences, 497: 38-55. https://doi.org/10.1016/j.ins.2019.05.035

[5] Safaa Mahdi, A., Mezaal Shati, N. (2024). A survey on fake news detection in social media using graph neural networks. Journal of Al-Qadisiyah for Computer Science and Mathematics, 16(2): Comp.23-41. https://doi.org/10.29304/jqcsm.2024.16.21539

[6] Yuan, L., Jiang, H., Shen, H., Shi, L., Cheng, N. (2023). Sustainable development of information dissemination: A review of current fake news detection research and practice. Systems, 11(9): 458. https://doi.org/10.3390/systems11090458

[7] Sanida, M.V., Sanida, T., Sideris, A., Dossis, M., Dasygenis, M. (2024). Fake news detection approach using hybrid deep learning framework. In 2024 9th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), Athens, Greece, pp. 81-84. https://doi.org/10.1109/SEEDA-CECNSM63478.2024.00023

[8] Huang, Y.F., Chen, P.H. (2020). Fake news detection using an ensemble learning model based on self-adaptive harmony search algorithms. Expert Systems with Applications, 159: 113584. https://doi.org/10.1016/j.eswa.2020.113584

[9] Khan, J.Y., Khondaker, M.T.I., Afroz, S., Uddin, G., Iqbal, A. (2021). A benchmark study of machine learning models for online fake news detection. Machine Learning with Applications, 4: 100032. https://doi.org/10.1016/j.mlwa.2021.100032

[10] Ghanem, B., Ponzetto, S.P., Rosso, P., Rangel, F. (2021). Fakeflow: Fake news detection by modeling the flow of affective information. arXiv preprint arXiv:2101.09810. https://doi.org/10.48550/arXiv.2101.09810

[11] Akinyemi, B., Adewusi, O., Oyebade, A. (2020). An improved classification model for fake news detection in social media. International Journal of Information Technology and Computer Science, 12(1): 34-43. https://doi.org/10.5815/ijitcs.2020.01.05

[12] Qawasmeh, E., Tawalbeh, M., Abdullah, M. (2019). Automatic identification of fake news using deep learning. In 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, pp. 383-388. https://doi.org/10.1109/SNAMS.2019.8931873

[13] Ahmed, H., Traore, I., Saad, S. (2017). Detection of online fake news using N-gram analysis and machine learning techniques. In Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, Springer, Cham. https://doi.org/10.1007/978-3-319-69155-8_9

[14] Albahar, M. (2021). A hybrid model for fake news detection: Leveraging news content and user comments in fake news. IET Information Security, 15(2): 169-177. https://doi.org/10.1049/ise2.12021

[15] Basri, M., Abd Rahim, N.H. (2022). Hybrid ensemble model for fake news detection. Journal of Theoretical and Applied Information Technology, 100(14): 5253-5262.

[16] Jardaneh, G., Abdelhaq, H., Buzz, M., Johnson, D. (2019). Classifying Arabic tweets based on credibility using content and user features. In 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, pp. 596-601. https://doi.org/10.1109/JEEIT.2019.8717386

[17] Abd, D.H., Mahdi, M.F., Jassim, M.A., Hussain, A. (2023). Arabic fake news detection using ensemble technique. In 2023 16th International Conference on Developments in eSystems Engineering (DeSE), Istanbul, Turkiye, pp. 292-297. https://doi.org/10.1109/DeSE60595.2023.10469046

[18] Alkhair, M., Meftouh, K., Smaïli, K., Othman, N. (2019). An Arabic corpus of fake news: Collection, analysis and classification. In Arabic Language Processing: From Theory to Practice, Springer, Cham. https://doi.org/10.1007/978-3-030-32959-4_21

[19] Hansrajh, A., Adeliyi, T.T., Wing, J. (2021). Detection of online fake news using blending ensemble learning. Scientific Programming, 2021(1): 3434458. https://doi.org/10.1155/2021/3434458

[20] Thaher, T., Saheb, M., Turabieh, H., Chantar, H. (2021). Intelligent detection of false information in Arabic tweets utilizing hybrid Harris Hawks based feature selection and machine learning models. Symmetry, 13(4): 556. https://doi.org/10.3390/sym13040556

[21] Alguttar, A.A., Shaaban, O.A., Yildirim, R. (2024). Optimized fake news classification: Leveraging ensembles learning and parameter tuning in machine and deep learning methods. Applied Artificial Intelligence, 38(1): 2385856. https://doi.org/10.1080/08839514.2024.2385856

[22] Alanazi, S.S., Khan, M.B. (2020). Arabic fake news detection in social media using readers' comments: Text mining techniques in action. International Journal of Computer Science and Network Security, 20(9): 29-35. https://doi.org/10.22937/IJCSNS.2020.20.09.4

[23] Lasotte, Y.B., Garba, E.J., Malgwi, Y.M., Buhari, M.A. (2022). An ensemble machine learning approach for fake news detection and classification using a soft voting classifier. European Journal of Electrical Engineering and Computer Science, 6(2): 1-7. https://doi.org/10.24018/ejece.2022.6.2.409

[24] Azizah, S.F.N., Cahyono, H.D., Sihwi, S.W., Widiarto, W. (2023). Performance analysis of transformer based models (BERT, ALBERT, and RoBERTa) in fake news detection. In 2023 6th International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, pp. 425-430. https://doi.org/10.1109/ICOIACT59844.2023.10455849

[25] Hermawan, A., Lunardi, L., Kurnia, Y., Daniawan, B., Junaedi, J. (2025). Optimizing convolutional neural networks with particle swarm optimization for enhanced

hoax news detection. Journal of Information Systems Engineering and Business Intelligence, 2(1): 53-64. https://doi.org/10.20473/jisebi.11.1.53-64

[26] Ahmed, H., Traore, I., Saad, S. (2017). Detection of online fake news using N-gram analysis and machine learning techniques. In Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, Springer, Cham. https://doi.org/10.1007/978-3-319-69155-8_9

[27] Emine, B. Fake News Detection Datasets, Kaggle. https://www.kaggle.com/datasets/emineyetm/fake-news-detection-datasets.

[28] Mehta, D., Patel, M., Dangi, A., Patwa, N., Patel, Z., Jain, R., Shah, P., Suthar, B. (2024). Exploring the efficacy of natural language processing and supervised learning in the classification of fake news articles. Advances of Robotic Technology, 2(1). https://doi.org/10.23880/art-16000108

[29] Abdalrdha, Z.K., Al-Bakry, A.M., Farhan, A.K. (2024). Crimes tweet detection based on CNN hyper parameter optimization using snake optimizer. In New Trends in Information and Communications Technology Applications, Springer, Cham. https://doi.org/10.1007/978-3-031-62814-6_15

[30] Zhang, H., Xiao, X., Mercaldo, F., Ni, S., Martinelli, F., Sangaiah, A.K. (2019). Classification of ransomware families with machine learning based on N-gram of opcodes. Future Generation Computer Systems, 90: 211-221. https://doi.org/10.1016/j.future.2018.07.052

[31] Jaleel, H.Q., Stephan, J.J., Naji, S.A. (2022). Textual dataset classification using supervised machine learning techniques. Engineering and Technology Journal, 40(4): 527-538. https://doi.org/10.30684/etj.v40i4.1970

[32] Dhall, D., Kaur, R., Juneja, M. (2020). Machine learning: A review of the algorithms and its applications. In Proceedings of ICRIC 2019, Springer, Cham. https://doi.org/10.1007/978-3-030-29407-6_5

[33] Ray, S., Srivastava, T., Dar, P., Shaikh, F. (2020). Understanding support vector machine algorithm from examples (along with code). https://www.analyticsvidhya.com/blog/2020/09/understaing-support-vector-machine-example-code/.

[34] Yuslee, N.S., Abdullah, N.A.S. (2021). Fake news detection using Naive Bayes. In 2021 IEEE 11th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, pp. 112-117. https://doi.org/10.1109/ICSET53708.2021.9612540

[35] Al-obaidi, S.A. (2024). Automated fake news detection system. Iraqi Journal for Computer Science and Mathematics, 5(4): 2. https://doi.org/10.52866/2788-7421.1200

[36] Chinta, S.V., Fernandes, K., Cheng, N., Fernandez, J., Yazdani, S., Yin, Z., Wang, Z., Wang, X., Xu, W., Liu, J., Yew, C.S., Jiang, P., Zhang, W. (2023). Optimization and improvement of fake news detection using voting technique for societal benefit. In 2023 IEEE International Conference on Data Mining Workshops (ICDMW), Shanghai, China, pp. 1565-1574. https://doi.org/10.1109/ICDMW60847.2023.00199

[37] Xie, H., Zhang, L., Lim, C.P., Yu, Y., Liu, H. (2021). Feature selection using enhanced particle swarm optimisation for classification models. Sensors, 21(5): 1816. https://doi.org/10.3390/s21051816

[38] Baratloo, A., Hosseini, M., Negida, A., El Ashal, G. (2015). Part 1: Simple definition and calculation of accuracy, sensitivity and specificity. Emergency, 3(2): 48-49.

[39] Saito, T., Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PloS One, 10(3): e0118432. https://doi.org/10.1371/journal.pone.0118432

[40] Tafvizi, A., Avci, B., Sundararajan, M. (2022). Attributing AUC-ROC to analyze binary classifier performance. arXiv preprint arXiv:2205.11781. https://doi.org/10.48550/arXiv.2205.11781